

ExBIN
PROJECT





Technický úvod do projektu univerzálního protokolu

<https://xbup.exbin.org>

Miroslav Hajda

Pro **OpenAlt 2021**

2021/11/06

O projektu

Projekt XBUP v kostce:

- Binární datový protokol / souborový formát
- Stromová struktura
- Vestavěné a uživatelské datové typy
- Transformační služba

Proč binárně?

- Univerzální pro všechny data
- Efektivnější reprezentace
- Komprese
- Šifrování, stránkování
- Jazyková nezávislost

Nevýhody binárního formátu

- Nejde snadno editovat v textovém editoru
- Složitější na použití
- Složitější na návrh
- Obtížné provést změny ve struktuře
- Data jsou hůře čitelná při částečném poškození

Existující textové formáty

ASN.1

SGML

HTML

CSS

XML

Textové konfigurace

CSV

JSON

YAML

Existující binární formáty

RIFF

ASN.1

HDF5

EXI (Binary XML)

BSON

UBJSON

EBML (Matroska)

Protocol Buffers

Thrift Protocol

CBOR

MessagePack

Cíle protokolu

- Univerzální
- Nezávislý
- Deklarativní
- Normativní
- Flexibilní
- Efektivní

Motivace pro použití protokolu

- Multimédia
- Serizalizace dat
- API pro aplikace
- Souborové archivy / systémy
- Vědecké účely a velká data
- Náhodný přístup a souběžné zpracování
- Archivace dat

Příklad formátu - obrázek

OBRÁZEK

PARAMETRY VZNIKU

ČAS VZNIKU

DIGITÁLNÍ PODPIS

GPS SOUŘADNICE

AUTOR

POUŽITÝ SOFTWARE

POZICE / POLOHA

SNÍMACÍ ZAŘÍZENÍ

ÚHEL POHLEDU

NÁZEV ZAŘÍZENÍ

TYP / ROZLIŠENÍ SENZORU

RYCHLOST POHYBU

VÝROBCE

DOBA OSVITU / UZÁVĚRKA

KOMENTÁŘ

OBRAZOVÁ DATA

ROZLIŠENÍ OBRÁZKU

KOMPRESSE

MODEL BAREV

BITMAPOVÁ DATA

PALETA

CRC / KONTROLNÍ SOUČET

Univerzální reprezentace dat

- Data jako sekvence bitů
- Organizace dat po bajtech
- Unární a binární kódování
- Minimalistický přístup - budování zdola nahoru
- Základní typ - nezáporné celé číslo

Dynamické kódování čísel

Délka čísla | vlastní hodnota
Unární kódování | binární kódování

0 0000000		= 0
0 0000001		= 1
0 0000010		= 2
0 0000011		= 3
...		
0 1111111		= 7Fh = 127
10 000000 00000000		= 80h = 128
10 000000 00000001		= 81h = 129
...		
10 111111 11111111		= 407Fh = 16511
110 00000 00000000 00000000		= 4080h = 16512
...		

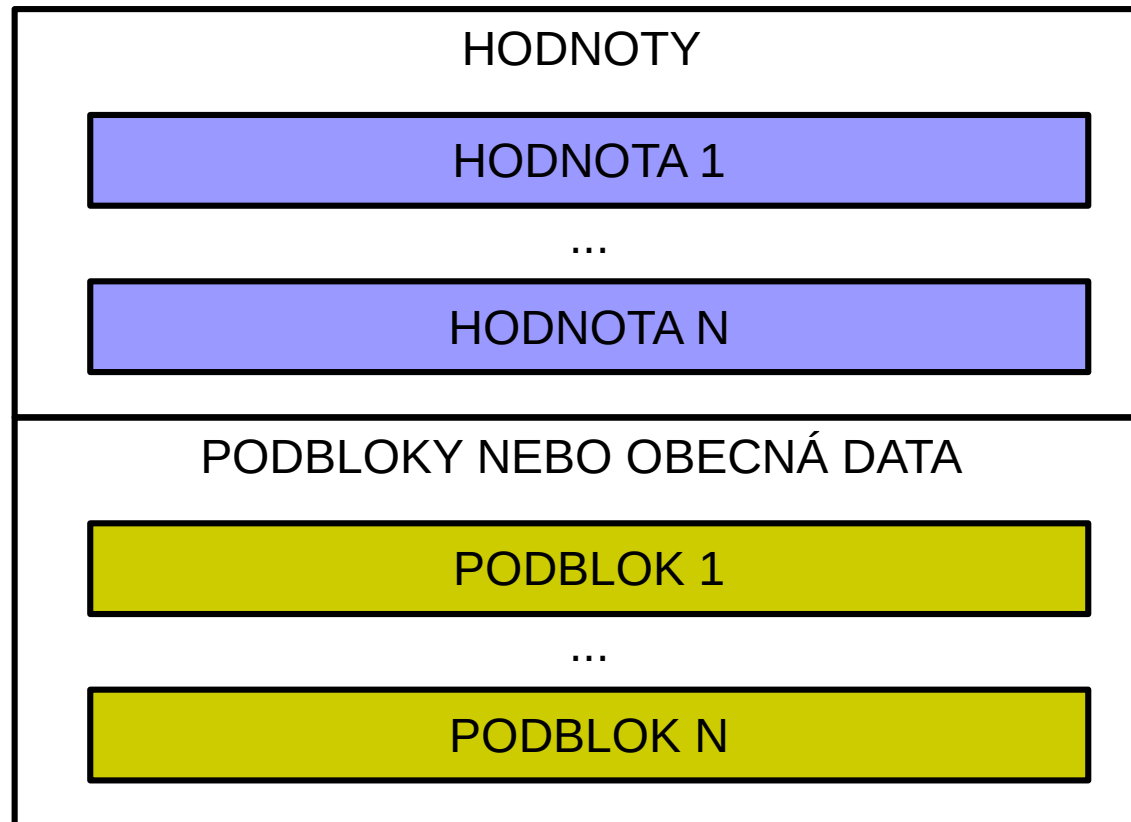
Blokově stromová struktura

- Univerzální stromová struktura
- Blok jako posloupnost hodnot + posloupnost podbloků
- Bloky lze odkazovat pomocí cesty ve stromu
- Podpora pro nekonečný proud dat

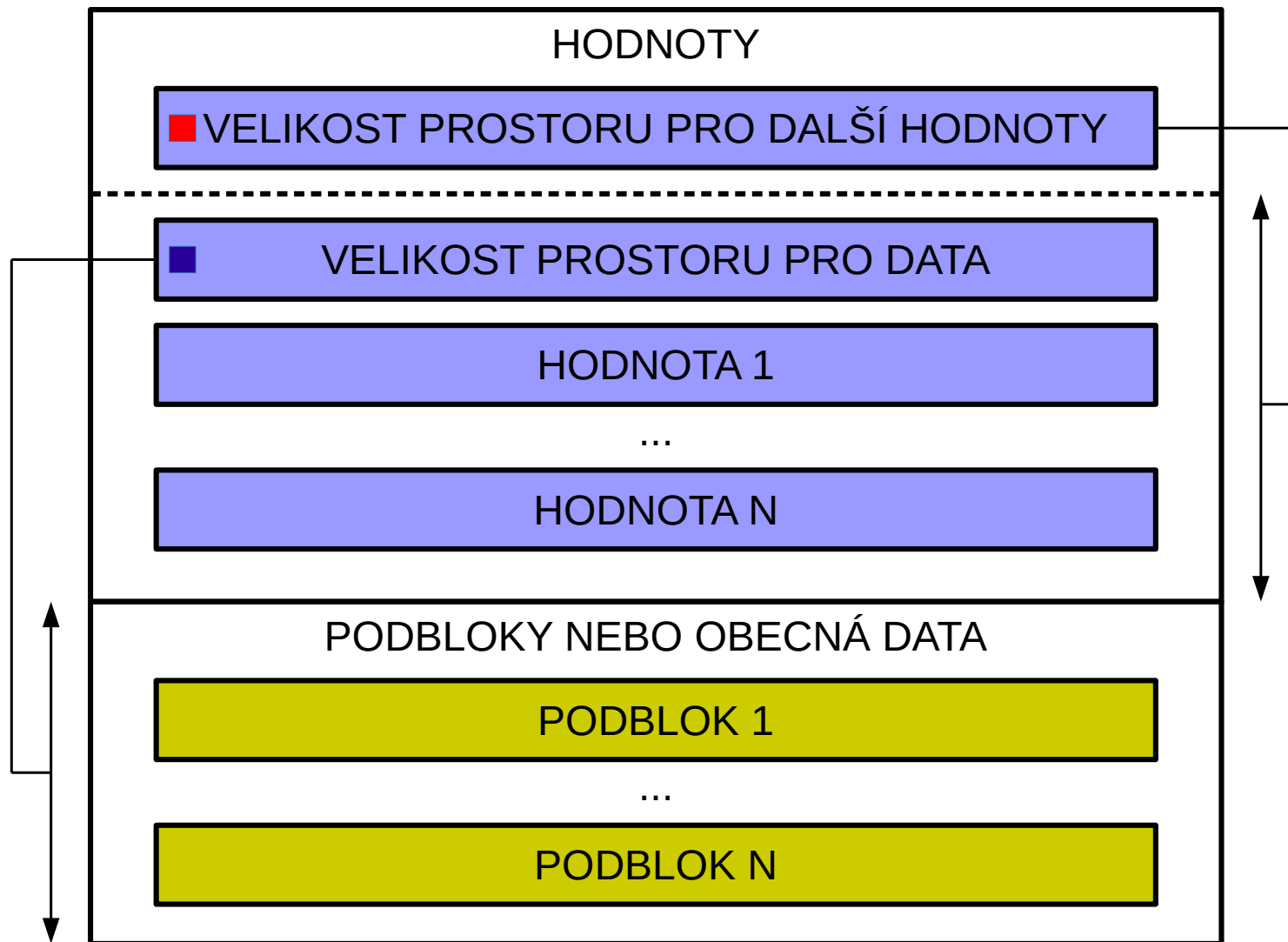
Struktura souboru



Datový blok



Datový blok detailněji



Příklady bloků

- [00] - Terminační blok (speciální význam)
- [01 00] - Prázdný blok bez dat
- [03 00 01 02] - Blok s hodnotami 1 a 2
- [01 02 12 34] - Blok s binárními daty [12 34]
- [02 02 05 01 00] - Blok s hodnotou 5 a jedním prázdným podblokem
- [02 7F 05 01 00 00] - Terminovaný blok
- [02 06 07 03 00 01 02 01 00]

Příklady bloků

- [00] - Terminační blok (speciální význam)
- [01 00] - Prázdný blok bez dat
- [03 00 01 02] - Blok s hodnotami 1 a 2
- [01 02 12 34] - Blok s binárními daty [12 34]
- [02 02 05 01 00] - Blok s hodnotou 5 a jedním prázdným podblokem
- [02 7F 05 01 00 00] - Terminovaný blok
- [02 06 07 03 00 01 02 01 00] - ?

Příklady bloků

- [00] - Terminační blok (speciální význam)
- [01 00] - Prázdný blok bez dat
- [03 00 01 02] - Blok s hodnotami 1 a 2
- [01 02 12 34] - Blok s binárními daty [12 34]
- [02 02 05 01 00] - Blok s hodnotou 5 a jedním prázdným podblokem
- [02 7F 05 01 00 00] - Terminovaný blok
- [02 06 07 03 00 01 02 01 00] - Blok s hodnotou 7 a dvěma podbloky

Definice typu dat

- První dvě hodnoty datového bloku určují typ
- Hlavní katalog pro základní typy dat
- Další typy dat lze definovat i externě
- Definice typu jako seznam položek (relační)
- Podpora pro podseznamy
- Typy se mohou spojovat a skládat dohromady

Příklad datového typu

OBRÁZEK

Rozlišení:

ROZMĚRY (X,Y) → 2 CELÁ ČÍSLA

Skutečný
rozměr:

ROZMĚRY (X,Y) → 2 REÁLNÉ ČÍSLA

↓
DÉLKA V METRECH

↓
JEDNOTKA : METR

↓
JEDNOTKA VZDÁLENOSTI

Matice pixelů:

3 ADDITIVNÍ BARVY (RGB)

↓
ČERVENÉ SVĚTLO

ZELENÉ SVĚTLO

MODRÉ SVĚTLO

→ FREKVENCE
ELMAG.
ZÁŘENÍ

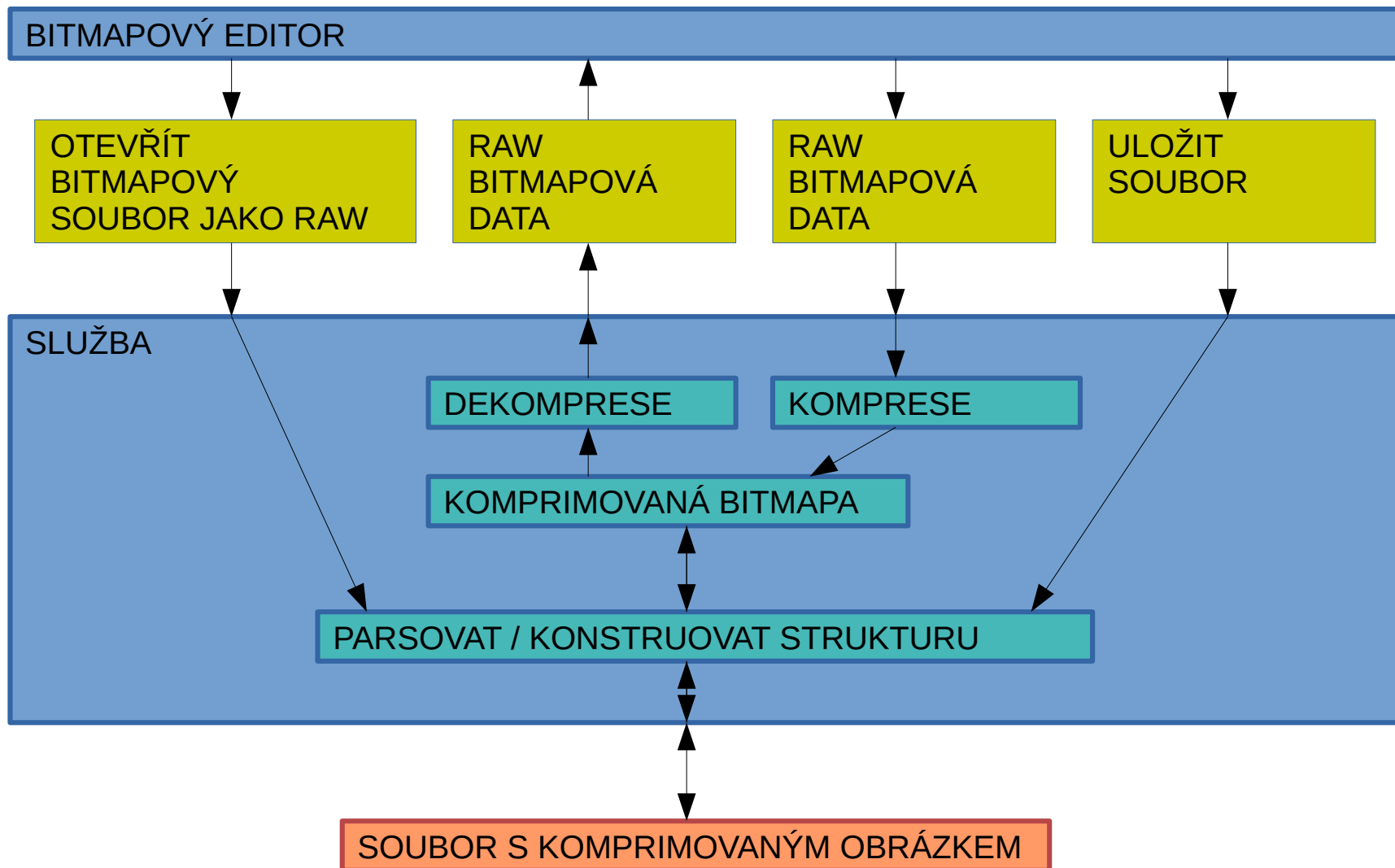
Katalog / Knihovna

- Položky katalogu jsou obecné entity
- Stromová struktura katalogu
- API pomocí protokolu nebo služby
- Vícejazykové názvy, ikony
- Rozšiřitelné o další typy entit

Transformace dat

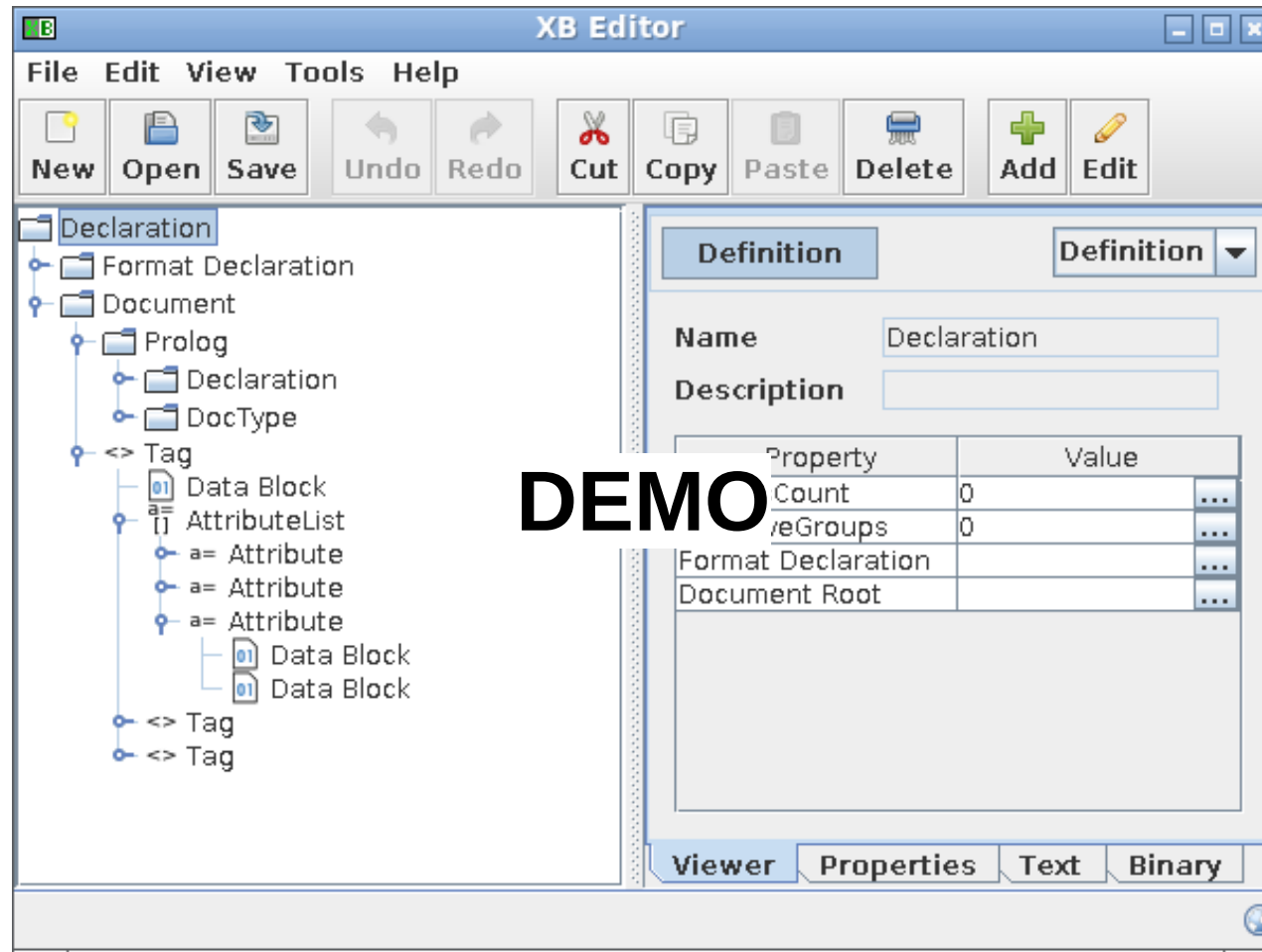
- Operace: vstupní data na výstupní data
- Je poskytováno jako služba OS / knihovna
- Služba automaticky převádí data oběma směry
- Kombinování transformací
- Řešení kompatibility

Příklad použití služby



Editor

- Zobrazení / editace stromové struktury
- Vizualizace dat (např. obrázky)
- Kde je definována syntaxe, je možné blok jako text
- Umožňuje podporu více syntaxí i jazyků
- Umožňuje provádět transformace (komprese...)



DEMO

Definition Definition ▾

Name Declaration

Description

Property	Value	
Count	0	...
reGroups	0	...
Format Declaration		...
Document Root		...

Viewer Properties Text Binary

Další kroky

Vytvořit:

- Specifikaci / RFC
- Referenční implementaci protokolu
- Katalog typů
- Referenční implementaci služby OS
- Ukázkovou aplikaci

Otázky?

- Web: <https://xbup.exbin.org/>
- Zdrojové kódy: <http://github.com/exbin>
- Element chat: [XBUP @ matrix.org](https://matrix.org)
- Email: exbinproject@gmail.com

Děkuji za zhlédnutí